

S/N 09/891, 080  
 ACT UNIT 2151

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 02-008976  
 (43)Date of publication of application : 12.01.1990

(51)Int.Cl. G06K 9/20  
 G06F 15/70

(21)Application number : 63-160065

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 28.06.1988

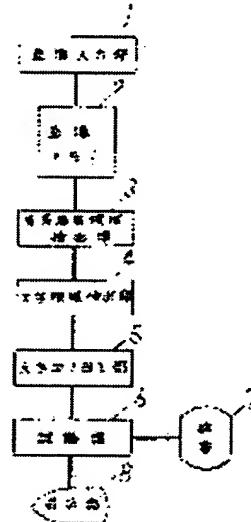
(72)Inventor : TAKENOUCHI MARIKO  
 FUJIWARA HIROE

## (54) CHARACTER RECOGNITION DEVICE

### (57)Abstract:

**PURPOSE:** To automatically decide an area targeted to be recognized by extracting a background image element area from an input image, and extracting a character area by checking mutual relation between the background image elements.

**CONSTITUTION:** An image inputted from an image input part 1 is stored in an image memory 2 as binary data in which the black image element of a character part is set at 1 and the white image element at 0. A background image element area extraction part 3 extracts a background image element area that is the separation candidate position of the character area in the area targeted to be recognized. And by checking the mutual relation between the background image elements at a character area extraction part 4, it is possible to automatically extract the character area that is the area targeted to be recognized.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

⑩ 日本国特許庁 (JP) ⑪ 特許出願公開  
⑫ 公開特許公報 (A) 平2-8976

⑬ Int. Cl. 5

G 06 K 9/20  
G 06 F 15/70

識別記号

3 4 0  
3 3 0

庁内整理番号

J 6942-5B  
Q 7368-5B

⑭ 公開 平成2年(1990)1月12日

審査請求 未請求 請求項の数 1, (全4頁)

⑮ 発明の名称 文字認識装置

⑯ 特願 昭63-160065

⑰ 出願 昭63(1988)6月28日

⑱ 発明者 竹之内 磨理子 大阪府門真市大字門真1006番地 松下電器産業株式会社内  
⑲ 発明者 藤原 啓恵 大阪府門真市大字門真1006番地 松下電器産業株式会社内  
⑳ 出願人 松下電器産業株式会社 大阪府門真市大字門真1006番地  
㉑ 代理人 弁理士 中尾 敏男 外1名

明細書

1、発明の名称

文字認識装置

2、特許請求の範囲

認識対象文字を含む画像を入力する画像入力部と、前記画像入力部で入力した画像から一定幅以上かつ一定高さ以上の背景画素領域を抽出する背景画素領域抽出部と、前記背景画素領域抽出部で抽出した背景画素領域相互の関係を調べて文字領域を抽出する文字領域抽出部と、前記文字領域抽出部で抽出した文字領域画像から認識対象文字を切り出す文字切り出し部と、前記文字切り出し部で切り出した認識対象文字を辞書と照合することにより認識候補文字を抽出する認識部を有するなどを特徴とする文字認識装置。

3、発明の詳細を説明

産業上の利用分野

本発明は、新聞、雑誌等の活字及び手書き文字を認識し、例えばJISコード等の情報量に変換する文字認識装置に関するものである。

従来の技術

従来の文字認識装置では、入力された画像を表示し、オペレータがマウス等を用いて設定した認識対象領域について認識を行っていた。また、他の文字認識装置では、画像入力部から入力する原画像自体にカラーペン等を用いてマークを付し、マークを認識することにより、自動的に認識対象領域を抽出していた(例えば、長谷他“マークを用いた印刷文書画像の領域抽出法”信学技報 Vol. 83 No. 302 PRL 83-78)。

発明が解決しようとする課題

しかしながら、オペレータやマークにより認識対象領域を指示する従来の技術では、例えば名刺のように空白部分で文字領域を分割し予め認識対象領域の候補領域が示されている文書に関しては、認識対象領域全域あるいは周囲を指定するという煩雑な作業を必要とした。

本発明はかかる点に鑑みてなされたものであり、既に空白部分で認識対象領域の候補領域が示されている文書に関して、簡易な方法で自動的に認識

対象領域を決定する文字認識装置を提供することを目的としている。

#### 課題を解決するための手段

本発明は前記課題を解決するため、画像入力部で入力した画像から一定幅以上かつ一定高さ以上の背景画素領域を抽出する背景画素領域抽出部と、背景画素領域抽出部で抽出した背景画素領域相互の関係を調べて文字領域を抽出する文字領域抽出部を備えている。

#### 作用

本発明は前記した技術手段により、背景画素領域抽出部で認識対象領域である文字領域の区切り候補位置である背景画素領域が抽出され、文字領域抽出部で背景画素領域相互の関係を調べることにより自動的に認識対象領域である文字領域を抽出することが可能となる。

#### 実施例

以下、本発明の実施例について図面を参照しながら説明する。

第1図は、本発明による文字認識装置の一実施

る文字領域の区切り候補位置である背景画素領域を抽出する。まず、画像メモリ2に蓄えられている入力画像Pを横方向に走査して、N画素以上(Nはあらかじめ設定)白画素が連続する横方向背景部ランを抽出する。第3図aに入力画像Pの横方向背景部ラン抽出結果を"—"で示す。同様に、画像メモリ2に蓄えられている入力画像Pを縦方向に走査して、N画素以上白画素が連続する縦方向背景部ランを抽出する。第3図bに入力画像Pの縦方向背景部ラン抽出結果を"|"で示す。次に、横方向背景部ランと縦方向背景部ランの両方に含まれる画素を背景画素とし、背景画素からなる背景画素領域を抽出する。第3図cに入力画像Pの背景画素領域抽出結果を示す。

文字領域抽出部4では、背景画素領域抽出部3で抽出した背景画素領域相互の関係を調べて認識対象領域である文字領域を抽出する。背景画素領域抽出部3で抽出した背景画素領域に含まれない画素の上下左右の連続性を調べて、連続している画素の繋りを文字領域とする。第4図に入力画像

例の構成図である。1は画像入力部であり、認識対象文字を含む画像を走査して2値信号で画像メモリ2に格納する。3は背景画素領域抽出部であり、文字領域の区切り候補位置である背景画素領域を抽出する。4は文字領域抽出部であり、背景画素領域抽出部3で抽出した背景画素領域相互の関係を調べて文字領域を抽出する。5は文字切り出し部であり、文字領域抽出部4で抽出した各々の文字領域毎に認識対象文字を1字ずつ切り出す。6は認識部であり、文字切り出し部5で切り出した認識対象文字の特徴量を求め、あらかじめ辞書7に登録されている文字の特徴量と照合し、最も似た文字を認識候補文字として表示部8に表示する。

以上のように構成された文字認識装置について、第2図に示す入力画像Pを例に説明する。

画像入力部1から入力された画像Pは文字部の黒画素を1、背景部の白画素を0の2値データで画像メモリ2に蓄えられる。

背景画素領域抽出部3では、認識対象領域であ

Pの文字領域抽出結果  $A_i$  ( $i = 1, 2, \dots, 6$ )を示す。

文字切り出し部5では、文字領域抽出部4で抽出した文字領域毎に認識対象文字を切り出す。第6図に入力画像Pの抽出された文字領域  $A_3$  を示す。まず、文字領域内の文字部全体の縦方向長さ  $D_v$  と横方向長さ  $D_h$  を比較し、文字列方向が縦書きか横書きかを判定する。文字領域  $A_3$  の場合、 $D_v > D_h$  より縦書きである。次に、文字列方向判定結果を用いて、文字領域を文字列方向と垂直に射影して文字部を形成する画素のヒストグラム  $H_c$  を求める。ヒストグラム  $H_c$  の値が1画素以上である文字部の開始・終了座標 ( $C_{s1}, C_{e1}$ ) …… ( $C_{s1}, C_{e1}$ ) ……を用いて認識対象文字画像を切り出す。

認識部6には切り出した認識対象文字画像Rが順次入力される。第8図aに文字領域  $A_3$  の切り出した認識対象文字『文』の画像を示す。認識部6では、切り出した『文』の各画素について、第8図bの矢印が示す方向に着目画素を含んでM個

以上(Mはあらかじめ設定)連っているか否かを調べ方向コードを設定する。方向コード毎に各画素の連結性を調べてストロークを抽出し、ストロークの数・位置・長さの特徴量を抽出する。第6図に認識対象文字『文』のストローク抽出結果を示す。抽出した特徴量を辞書7に登録されている文字の特徴量と照合し、最も似た文字「文」を認識候補文字とし、表示部8に表示する。

#### 発明の効果

本発明によれば、入力画像から一定幅以上かつ一定高さ以上の背景画素領域を抽出し、背景画素領域相互の関係を調べて文字領域を抽出することにより、既に空白部分で認識対象領域の候補領域が示されている文書に関して、簡易な方法で自動的に認識対象領域を決定することが可能となる。

#### 4. 図面の簡単な説明

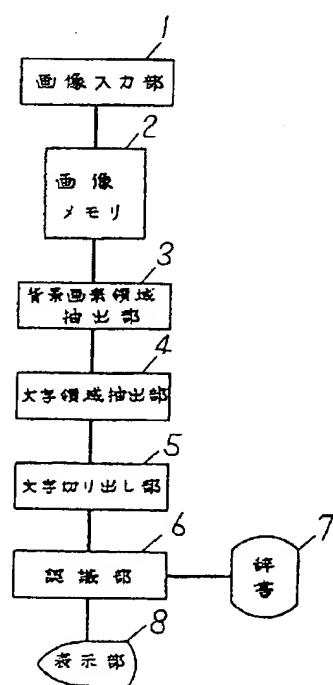
第1図は本発明における一実施例による文字認識装置の構成図、第2図は入力画像の説明図、第3図は背景画素領域抽出方法の説明図、第4図は文字領域抽出方法の説明図、第5図は文字切り出

し方法の説明図、第6図は文字認識方法の説明図である。

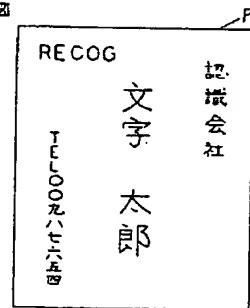
3……背景画素領域抽出部、4……文字領域抽出部、5……文字切り出し部、6……認識部。

代理人の氏名 弁理士 中尾敏男 ほか1名

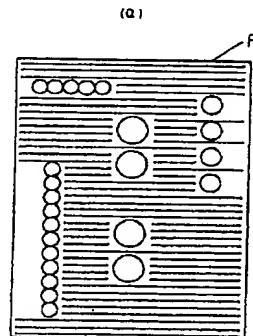
第1図



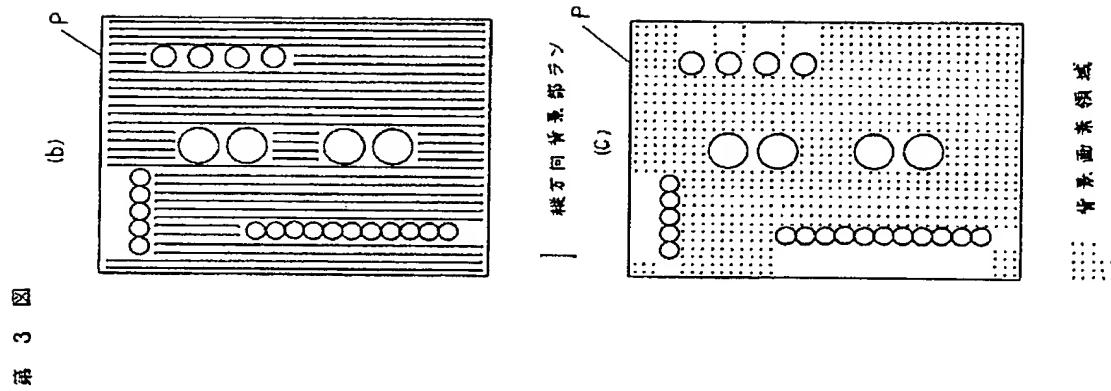
第2図



第3図

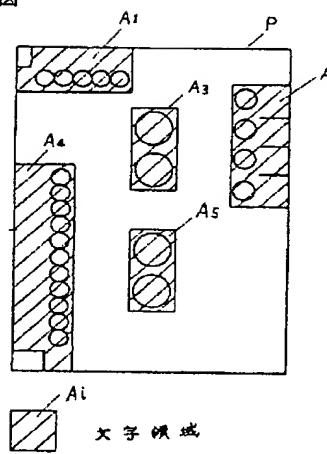


— 横方向背景部ラン

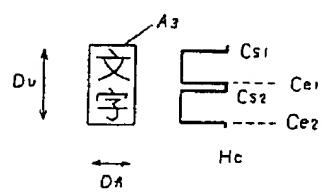


第 3 図

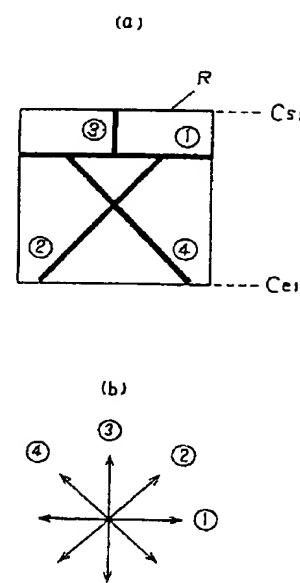
第 4 図



第 5 図



第 6 図



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.